CHAPTER 1

The Strangeness of Functions

Why do zebras have stripes? Biologists have argued about this since at least Darwin's time. Darwin himself dismissed the popular view that the stripes' purpose is camouflage: "The zebra is conspicuously striped, and stripes on the open plains of South Africa cannot afford any protection" (1871, 302). Others insist that stripes aren't there for camouflage but for cooling the animal (Larison et al. 2015). They think the black-and-white pattern chills the air around it. A third idea is that stripes play a role in social cohesion; the striped pattern draws zebras together into herds (Macdonald 2009, 689). A fourth possibility is that zebra stripes have no function at all (although I don't know of anyone who argues this in the literature). Maybe they're as biologically pointless as birthmarks, freckles, and chin clefts.

Recently, an American biologist, Tim Caro, threw his weight behind a newer idea (Caro et al. 2014; but see Harris 1930). He thinks the stripes' purpose is to deter biting flies. One particular family, the glossinids (commonly known as tsetse flies), is particularly troublesome, since it harbors a parasite responsible for African trypanosomiasis – the infamous sleeping sickness. Field and laboratory studies suggest that tsetse flies and other biting flies are averse to striped surfaces. Perhaps zebras use stripes to exploit this neurological quirk of the tsetse fly. Caro's hypothesis about the stripes' function is based on a mix of historical, geographical, and laboratory evidence, although the whole subject remains mired in controversy.

The parts and processes of the tsetse fly have functions, too. The tsetse fly is a family of bloodsucking flies that inhabit Central Africa, from the Sahara in the north to the Kalahari in the south. Unlike ordinary houseflies, it has a long, hollow proboscis. The tip of the proboscis is lined with tiny, sharp teeth, like a knife's serrated edge (see Krenn and Aspöck 2012, 111, Figure 8). The fly repeatedly prods an animal's thick hide until it draws blood. Its pharynx functions as a pump that sucks up the nutritious broth. A second pump shoots saliva into the wound in order to stop the blood from coagulating. The trypanosome parasite, *T. brucei*, lives in the saliva.

The parasite *T. brucei* has functions of its very own. It is one of many unicellular species of African trypanosome, and it resembles a tiny seahorse. It is the parasite responsible for sleeping sickness. Its coat contains millions of proteins called *variant surface glycoproteins*. The function of these proteins is to help *T. brucei* evade detection by the host's immune system (Horn and McCulloch 2010). The coat's genetic makeup is constantly changing: By the time the host's immune system learns to recognize one coat, *T. brucei* has morphed into another. As one geneticist described it to me, it is like changing hats, and the parasite changes its hat about once a week.

Functions are ubiquitous in the living world. Sometimes they harmonize; sometimes they clash. What are functions? At first glance, functions seem easy to understand. If functions are easy to understand, we should be able to give a clear and satisfying account of what they are. Instead, we find puzzles, and even contradictions, that drive us deeper into the nature of the living world.

When I ask biologists what functions are, I often get a similar response: "A trait's function is just what it does." Sometimes these biologists seem perplexed, and even mildly annoyed, to be asked a question like that. Hearts pump blood. That is what they do, so that is their function. Zebra stripes deter flies. That is what *they* do, so that is their function. The tsetse flies' labellar teeth puncture skin; *T. brucei's* glycoprotein coat tricks the host's immune system. Functions are simply doings.

Sadly, the biologists' simple account can't be right – for two reasons. First, traits do many things that aren't their functions. Noses help us breathe; they also hold up glasses, but their function is to help us breathe, not hold up glasses. Holding up glasses is a lucky benefit, or side effect, but not a function. Zebra stripes entertain safari guests, but that's also not their function. To use philosophical lingo, the fact that stripes entertain safarigoers is an "accident" and not a function. A good account of function should help us understand how functions and accidents differ.

Here's a second problem with the simple account that says a trait's function is just what it does. A particular instance of a trait – my stomach, your heart – can have the function of doing something even if it can't actually do that thing. If my stomach shuts down because of a drug overdose, it can't digest food. Yet it has the function of digesting food (it's a stomach, after all); thus it has a function it can't perform. It's "dysfunctional" or, if you prefer, "malfunctioning." Philosophers sometimes call this feature of functions – that it's possible for a trait to have a function it cannot perform, that it can dysfunction or malfunction – the "normativity" of functions. A good theory of function should make sense of this normativity, too.

We have to be careful here. Scientists use the word "function" in different ways; one theory won't fit all uses. I need to home in on the sense I'm after. In one sense of "function," functions are just *effects*. Climate change is a function of deforestation. Poor academic performance is a function of malnutrition. That isn't the sense of "function" I want to know about, and it's not the one that's prominent in biology. That sense of "function" doesn't let us distinguish functions and accidents; nor does it have a normative side. Whenever I use the term "function" without qualification, I mean it in the ordinary biological sense, where functions differ from accidents and where things sometimes malfunction. In Chapter 9, I'll return to the problem of how different senses of "function" fit together.

Here's the plan for this chapter. In Section 1.1, I'll turn to ordinary biological usage to extract a vital clue about functions: namely, functions have *explanatory depth*. When biologists give functions to traits, they often purport to give causal explanations for why those traits exist. By meditating on this one feature of functions, we can solve the other main puzzles of function. The problem is that, at first glance, it's hard to see how functions can actually explain anything; this is the so-called problem of *backwards causation* (Section 1.2). Section 1.3 will survey some of the more adventurous ways people have tried to solve backwards causation. It will also introduce the idea that, to solve backwards causation, functions should be selected effects. In Section 1.4, I'll be a bit more precise about what explanatory depth amounts to: that is, what functions are *supposed* to explain, whether or not they actually do so. In Section 1.5, I'll set out the ground rules of the game: What exactly is a theory of function, and what kinds of evidence should one draw on to support such a theory?

1.1 Functions and Explanations

What are functions? One clue comes from considering their role in explanations, in the practice of asking and responding to why-questions in science and in everyday life. Sometimes, when people say that a trait has a function, they're trying to explain why that trait exists. For example, sometimes, when scientists wonder about the function of zebra stripes, they're just wondering why zebras have stripes (rather than, say, being monocolored like some horses). And sometimes, when scientists argue with each another about the stripes' function, they're just arguing about how zebras came to have stripes. Maybe if we think about how functions fit into explanations, we'll come closer to understanding these other puzzling features of function. (In Section 1.5, I'll defend my method of figuring out what functions are: Examine how biologists use the term "function," and then step back and figure out what functions must actually be to support their usage.)

It's helpful to have some concrete examples in front of us. Tim Caro and his colleagues (2014), whom I alluded to earlier, wrote a paper simply entitled "The function of zebra stripes." It appeared in a major scientific journal, *Nature Communications*. From the outset, Caro makes it clear that solving the riddle of the stripes' function *just amounts to* explaining why zebras have stripes. Three pieces of textual evidence back up this interpretation. First, Caro writes of five different "functional hypotheses" about zebra stripes; he also calls these "factors proposed for driving the evolution of zebras' extraordinary coat coloration" (2). For Caro, offering a "functional hypothesis" about stripes and making a conjecture about why stripes evolved are one and the same thing. It's not that there are two questions, one about functions and one about origins. There's one question that can be posed using different words.

Second, to support his hypothesis about the stripes' function, Caro collected historical data about the distribution of tsetse flies in the regions he studied, as well as the historical distribution of various predators. Given that he wanted to show why stripes evolved, this was a sensible thing to do. Tsetse flies must have actually lived in the zebra's habitat when stripes evolved; otherwise, his "functional hypothesis" would be demolished. If Caro didn't care about why stripes evolved, why bother collecting historical data, which can be quite time-consuming and labor-intensive?

One final fact about Caro's paper seems noteworthy. A few years later, he and his colleagues wrote another paper with the elegant title, "Why is the giant panda black and white?" This one appeared in the journal *Behavioral Ecology*. It's plausible to think that the question he was asking about zebra stripes is the exact same kind of question he was asking about the panda's unique constellation of markings. The difference is verbal: One title is phrased in terms of functions, and the other is phrased in terms of why a trait exists. My point isn't that Caro is right about the function of stripes. My point is that, in some cases, function statements purport to be explanations. They have explanatory depth.

Here's another example of how biologists fuse functions and explanations. In his book, *Neural Activity and the Growth of the Brain* (1994), the neuroscientist Dale Purves discusses how hard it is to decipher the function of animal markings. I'll cite a long passage because it reveals the explanatory depth of functions:

> Skin and fur markings are so striking that it is natural to assume that they must reflect some fundamental *function* of the integument. The major *purposes* of the skin, however, are temperature control, water regulation, and protection from infection. In fact, zoologists have often found it rather hard to decipher the *role* of particular animal markings. Such patterns are sometimes used for camouflage or sexual attraction, but more often than not it is difficult to say just *why they are there*... (30; my emphases)

What is so striking about the passage is that Purves uses the expressions, "purpose," "function," "role," and "why they are there," interchangeably. In his usage, to state something's function just amounts to saying why it's there. Purves doesn't go so far as to say that all markings have functions. Maybe some markings, like birthmarks or freckles, are purposeless. My point is that, sometimes, when scientists give functions to traits, they're trying to explain why those traits exist.

Here's a final example of how biologists tie functions to explanations. The Harvard molecular biologist Sean Eddy, in a discussion of the concept of "nonfunctional DNA," tells us that "by nonfunctional, we mean 'having little or no selective advantage for the organism'" (2012, R898). This makes it sound as if the function of a stretch of DNA is simply whatever it does that promotes the organism's relative fitness, regardless of how it got there. Later, however, he clarifies that when we argue about whether stretches of DNA are functional, we're arguing about "whether *they're there* primarily because they're useful for the organism (R898, emphasis mine)." So, for Eddy, when we say something's functional, we're not saying (or not only saying) that it does something or other to help us out, but that the fact that it helps us out in the specified way is the reason it exists.

Scientists aren't alone in using "function" with explanatory depth. Laypeople do, too. Several newspapers reported Caro's work on the function of stripes, and they used expressions like "why zebras have stripes," or "why stripes evolved," synonymously with "the function of stripes." Everyone seemed to understand that Caro's paper, "The function of zebra stripes," was about why zebras have stripes. The idea that functions are explanations is not some philosopher's invention. It's a robust feature of how scientists and laypeople alike think and talk about them.

This is a remarkable feature of functions, and one we shouldn't brush under the rug. In fact, philosophers of science have puzzled over the explanatory depth of functions ever since they started thinking seriously about them. I trace the modern functions debate back to the philosophers Carl Hempel (1965) and Ernest Nagel (1953). They agreed that function statements often *purport* to be explanations; they took this explanatory ambition as a plain fact of ordinary biological talk. They disagreed, however, about whether functions *actually* explain anything: that is, whether this explanatory ambition is ever fulfilled. Before we get tied up in debates, though, we should nail down exactly what is at issue. If function statements are supposed to be explanations, what is supposed to explain what? And what kind of explanation is on offer?

Consider how the popular press reported Tim Caro's work. According to the journalists, Caro and his colleagues discovered why zebras have stripes. The reason zebras have stripes is that stripes deter biting flies. It appears that, in Caro's functional hypothesis, the fact that stripes deter biting flies is supposed to explain why zebras have stripes.

To use philosophical lingo again, an explanation splits into two halves: the *explanandum* and the *explanans*. The explanandum is the fact or event one wishes to explain: why zebras have stripes; why Booth shot Lincoln; why rocks fall to the ground when you let them go. And the explanans is the fact or event that does the explaining: stripes deter flies; Booth wanted to extend the war; space-time is curved. If the statement, "The function of zebra stripes is to deter biting flies," is an explanation, the explanans is: Stripes deter flies. The explanandum is: Zebras have stripes.

Explanans: Stripes deter flies
explains
Explanandum: Zebras have stripes

Assuming that functions are explanatory, what kind of explanation is on offer? Explanations fall into different categories. There are, for instance, mathematical explanations, causal explanations, reductionist explanations, and statistical explanations (see Salmon 1989). The most natural answer here – the one that best fits the examples above – is that function statements are *causal* explanations, since they have to do with how things came to be. In a typical causal explanation, one explains why one event happens by pointing to an event that came before it. Why did Selena make the dean's list? Because she worked very hard in school. Why is there a coffee stain on the carpet? Because I knocked the mug over when I was reaching for the corn flakes. (Later, I'll consider, and reject, the idea that functions are explanatory in some noncausal sense.)

If function statements are explanations of the causal sort, then when we say the function of the zebra's stripes is to deter biting flies, we're saying that the fact that stripes deter flies causes zebras to have stripes. At any rate, that's the most natural way of interpreting biologists like Caro, Purves, and Eddy:

> Explanans: Stripes deter flies
> *causes*
> Explanandum: Zebras have stripes

From this perspective, when scientists give functions to traits, they're just giving extremely compact causal explanations for why those traits exist. Any theory of function that takes the causal-explanatory role of functions seriously is called an *etiological* approach to function, because "etiology" refers to the study of causes. The etiological approach to function isn't a single theory. It's a family of theories joined by the idea that function statements are causal explanations for the existence of traits. Its *locus classicus* is Wright (1973) – although as we'll see, Wright's specific version was somewhat off the mark.

This way of thinking about functions, where functions are just condensed causal explanations, is very attractive. In addition to reflecting the way biologists think and talk, it illuminates the two puzzling features of function we started with. (I'll come back to this point in the next chapter.) First, it shows how functions differ from accidents. The reason a function of the nose is to help us breathe, and not hold up glasses, is because the fact that noses help us breathe explains why we have them. The fact that noses are good at holding up glasses isn't why noses are there. The etiological approach also makes sense of function's "normativity" – that is, the possibility of malfunction or dysfunction. Something dysfunctions when it cannot do the thing that explains why it's there.

As an aside, when I say functions are "normative," I don't mean anything very nuanced or sophisticated. All I mean is that it's possible for something to dysfunction. Once we've explained how it's possible for something to dysfunction (see Chapter 8), we've explained function's "normativity." In my usage, there's nothing else hiding behind the word, nothing having to do with values or goals, oughts and shoulds, prescriptions or commands, the good or the just. Sometimes when I explain to people how dysfunctions are possible, they say things like, "Yes, I see, but how do you explain function's *normativity*?" In my preferred usage, there is no additional question here, but I don't wish to legislate usage for everyone else.

1.2 Backwards Causation

Despite its merits, etiological approaches to function seem to suffer a major drawback. This is known as the problem of "backwards causation" (e.g., Ruse 1973, 176). On the face of it, the fact that stripes deter flies cannot possibly explain why zebras have stripes. In a standard causal explanation, the relationship between cause and effect is a before-and-after one. In order for my knocking over the coffee cup to cause a carpet stain, I first knock over the cup, and then the stain appears, and not the other way around. (There might be exceptions to this rule. Kant noted that when a ball sits on a cushion and causes an impression on it, the ball and the cushion exist simultaneously. I won't linger on this because it isn't relevant to the typical biological examples.)

In order for the fact that stripes deter flies to cause zebras to have stripes, the two facts would then have to stand in a before-and-after relationship. It would have to be the case that, first, zebra stripes deter flies and, second, zebras have stripes, but that's the reverse of what happens in the real world. For a zebra to use its stripes to deter flies, it must already have the stripes. Hence the "backwards" part of the problem of backwards causation. In order for the fact that stripes deter flies to cause zebras to have stripes, there would have to be a peculiar causal relationship in which later events (a zebra using its stripes to deter flies) cause earlier ones (zebras have stripes). The etiological approach to function seems to flip the normal chronological order of causation. It violates a cornerstone of our understanding of cause and effect, at least as it applies to things outside of the weird quantum realm.

Here's a simple, no-nonsense solution to backwards causation: Flatly deny that function statements are causal explanations. This is what Hempel (1965), in effect, did. In fact, he denied that they're explanations at all. As he put it, "the information typically provided by a functional analysis of an item *i* affords neither deductively nor inductively adequate grounds for expecting *i* rather than one of its alternatives" (p. 313). In short, function statements don't actually explain why the functional item exists. Hempel admitted that scientists often think that functions are explanatory. He just thought they were in the grip of a cognitive illusion: "The impression that a functional analysis. . .explains the occurrence of *i*, is no doubt at least partly due to the benefit of hindsight: when we seek to explain an item *i*, we presumably know already that *i* has occurred" (ibid.).

Many philosophers have chosen to follow Hempel in denying that functions statements are causal explanations. They have worked out a vast array of *nonetiological* accounts of function instead. Fitness-contribution theories, such as the propensity theory and the biostatistical theory, hold that a trait's function has to do with its present-day contribution to fitness. Causal role theories say it has to do with the contribution a part makes, in tandem with other parts, to an interesting system capacity. Modal theories say the function of a trait has to do with its behavior on nearby possible worlds. Some versions of the organizational theory hold that a trait's function has to do with how it contributes to the persistence of that very trait. (Most of these will come in for scrutiny at various points of the book.) I have mapped out those positions in detail elsewhere, and outlined their relative strengths and weaknesses (Garson 2016). The thread that joins them together is their shared rejection of the idea that functions are causal explanations for traits.

I think it would be unfortunate if functions never provided successful causal explanations for traits, since so many scientists, science writers, and journalists, think they do. How could they all be so wrong? The fact that it would be unfortunate, however, shouldn't count for much. Sometimes, philosophy reveals that our deeply held assumptions about language, ethics, or reality, are mistaken; maybe this is just one of those occasions. Let's not give up too quickly though. We can at least try to find a coherent theory of function that shows how functions can be successful causal explanations. If our best efforts in this direction repeatedly end in frustration, we should be ready to embrace a *nonetiological* theory instead.

1.3 Theism and Fictionalism

Before throwing out the etiological approach because of backwards causation, here's a thought that should give us pause. At least sometimes, the effect of an item does explain "why it is there" – that is, why it is where it is. Suppose I ingest some beneficial bacteria (such as *Lactobacillus*) to promote good digestion, because I know it will raise gut acidity. In that case, the *Lactobacillus* is there, in my gut, *because* it promotes acidity. There's no mystical or mysterious backwards causation taking place. The fact that bacteria are good at promoting acidity caused me to put them in my gut.

Consider a somewhat more extreme case. Suppose we rescue a species from extinction because of some benefit it provides. In the early 1990s, southern California zoos decided to rescue California condors (*Gymnogyps californianus*) from extinction because they are huge, magnificent creatures, and they are good scavengers, too. The reason condors exist – exist *now*, *today* – is because they are huge, magnificent, and good at scavenging. If they possessed none of those features, we would no longer enjoy their presence. There's no mysterious backwards causation taking place when we say that condors exist because they are huge, magnificent scavengers.

Unfortunately, this solution to backwards causation doesn't apply to the ordinary biological cases, like the zebra's stripes, the tsetse fly's labellar teeth, or the ever-morphing coat of *T. brucei*. We solved backwards causation by bringing in intelligence and mind – that is, the intelligence of human beings and their mental states like beliefs and desires. The fact that *Lactobacillus* promotes gut acidity caused my *belief* that it promotes gut acidity; my *belief*, combined with my *desire* for a healthy digestive system, caused me to put them in my stomach. At best, this solution only applies to intelligent beings, and the things we do or make.

Of course, if we were willing to embrace theism, the problem of backwards causation would disappear. Perhaps there is a God who freely shapes the biological world, just as I freely shape my gut biota. Maybe God knew that zebra stripes would deter flies, and God gave stripes to the zebra for that reason, out of the kindness that defines God's very nature. Were that the case, we could affirm that zebras have stripes because stripes deter flies. (For that matter, perhaps the world as we know it is a mad scientist's computer simulation, and this scientist gave the zebra stripes just to confuse and bewilder our brightest minds. In that case, the stripes would still have a function, though one very different than we imagined.)

Considerations such as these led some philosophers to an adventurous position. They think that if there are any biological functions in the natural world, an intelligent being must have put them there. In other words, some consider it a conceptual truth that function requires intentional design. One can't accept functions, but reject God, in one and the same breath.

The philosopher Alvin Plantinga (1993, 214) deploys this idea as part of an innovative argument for God's existence (though one that echoes Aquinas, as he freely admits). His first premise is that there are functions in the natural world, as any biologist will tell you. (I invite you to track some down and ask them yourself.) His second premise is that, as a conceptual truth, functions require design. Thus, God is real, and a millennia-old controversy is solved.

Michael Ruse (2002), like Plantinga, thinks function requires design, but he reaches an even more adventurous conclusion. He agrees that, as a point of definition, if anything has a function, it must have been designed. Biologists talk about functions, he says, "because organisms. . .are taken to be design-like: they are taken to be as if they were artefacts, or parts of artefacts, created by conscious intelligences in order to serve certain ends" (2002, 37). Instead of embracing God, he denies functions. Strictly speaking, functions don't exist. When biologists talk about functions, they are trading in metaphors; biologists are playing a game of make-believe (p. 40). However – and this is the strange part – Ruse thinks biologists shouldn't stop playing this game of make-believe. That's because functions have "key heuristic value"(p. 46). When biologists look at organisms as if they were designed, they often discover new things about them. Philosophers often call this sort of view "fictionalism."

I want to avoid both of these extremes, theism and fictionalism. I'd like to think that functions exist, just like biologists think they do. The idea that science should embrace known falsehoods as an engine of discovery runs against its core mandate. I'd also like to think one could acknowledge that functions are real without taking a position on the existence of God. Finally, I want to acknowledge the explanatory depth of functions – that is, function statements are, sometimes, correct causal explanations for traits, as biologists like Caro think they are. What to do?

Fortunately, there is a solution. We can solve the problem of backwards causation by appealing to selection processes, not design. In the next chapter, I'll lay out the core argument for the traditional selected effects theory. This theory, in its most unadorned version, says that a trait's function is what it was selected for. I'll also defend the integrity of that solution from common complaints. The crucial point here is that the reason selected effects theorists tie functions to selection is neither because they think natural selection is the only force of evolutionary change, nor because they think natural selection is itself a source of design, nor because they don't understand the history of biology. (All of these are accusations that philosophers have leveled against selected effects theorists.) Rather, the selected effects theory is the best way to solve a conceptual problem that has plagued the functions debate for over sixty years: How do functions explain anything?

1.4 Being There

A few more aspects of function deserve clarification before moving on. Functions purport to explain why a trait "is there." Can we clarify this "being there"? What exactly are function statements supposed to explain? To be precise, when we say that some activity of a trait is its function, we purport to give a causal explanation. The explanandum – the thing we're trying to explain – can take two forms. Sometimes, we want to explain a fact about some particular entity: namely, why that entity has that trait. (Why does my rat Gemini stand on his hind legs when he wants some celery?) Other times, we want to explain a fact about a population. We want to explain why some, or most, or all members of a population have a trait. (Why do people have noses? Why do zebras have stripes?) Some, like Buller (1998, 522), think function statements are primarily trying to explain facts about individuals (why Gemini stands on his hind legs). Others, like Neander, (1991, 174), think functions statements are primarily trying to explain facts about populations (why people have noses). I don't see any good reason for restricting the explanandum of function statements either way. Sometimes, function statements are about individuals, and sometimes they're about populations.

Another question that comes up, by way of clarifying the explanandum, is this: What is it for an individual or group to "have" a trait? To say that an entity (individual or group) has a trait is to say that the trait is a physical, or psychological, or behavioral feature of that entity. (I'd use the term "part," but that's a bit too restrictive since it connotes physical parts. "Feature" is broad enough to capture behavioral or psychological characteristics, too.) Crucially, for an item to have a function, it must be a feature of a system, like an organism, cell, or group; but the system, considered as a whole, has no function. Hence, we can ask about the function of the zebra's stripes, or the function of the prairie moles' grooming behavior, or whether depression has a function in mammals. We can wonder about the function of group traits, like V-formation in a flock of geese, or predator signaling in a vervet colony. But the prairie mole itself has no function; nor does the zebra; nor does the flock.

This restriction – functions belong to features of systems but not to the system itself – is enshrined in ordinary biological usage, just like the claim that functions have explanatory depth. To see this, one need only consider the following thought experiment. When Caro and his colleagues wrote a paper entitled "The function of zebra stripes," the title was easy to understand. Had they written a paper called "The function of *zebras*," nobody would have known what they meant. Similarly, when he wrote, "Why is the giant panda black and white?," everyone knew what he was getting at. Had he written, instead, "Why are there pandas?," it would not have been published by a major scientific journal. (In the tradition of natural theology, which reached its apex with William Paley's 1802 book by the same title, one could still pose questions in this matter; e.g., why do mosquitoes exist? Biologists stopped asking questions like that by the middle of the nineteenth century for reasons that need not be rehearsed here.)

This way of thinking about functions, where functions belong, first and foremost, to *features* of individuals or groups, has a drawback. It prevents certain things from having functions that we might think of as having them. Consider a beaver dam or a honeycomb (Griffiths 1993, 416). Surely, the honeycomb has a function for the bee population: namely, to contain larva and store honey, but the honeycomb is not a feature of any particular bee or even a feature of a whole population of bees. According to my way of thinking, the honeycomb does not have a function, or at least not in the way that the bee's stinger has a function, but this is an acceptable limitation. We can say everything we'd like to say about functions without giving functions to honeycombs and beaver dams. Instead of talking about the function of the honeycomb, we can talk about the function of the comb-making behavior, which is a feature of the bee. There's no need to insist that honeycombs or beaver dams have biological functions of their very own. (As one reviewer pointed out, we might also solve this by considering the honeycomb as part of the bee's *extended phenotype*, in which case we could talk about the function of the honeycomb just as we talk about the function of the comb-making behavior – see Dawkins 1982. I won't pursue this suggestion here.)

Now, beaver dams and honeycombs do have "functions," after a certain manner. A honeycomb has an *artifact* function, much like the components of my old flip phone, or a car's engine, or clay pottery from the Neolithic era. As I'll discuss in the next chapter, artifact functions aren't biological functions, and we shouldn't force them into the same mold. We need one theory for biological functions, and another theory for artifact functions. It seems to me that beaver dams acquire their functions, in the first place, because individuals design, create, and use them; they don't get their functions in the same way that features of organisms do.

I realize that some entities have both biological and artifact functions, such as an edited segment of DNA or a marijuana-sniffing hound. This complicates the picture, but it doesn't change the fact that biological and artifact functions are different kinds of things. Why do I think this? Because the best theory of function on the market is the selected effects theory (to be precise, a particular version called the generalized selected effects theory) and that theory implies that, as a rule, artifact functions aren't biological functions, as I'll show in the next chapter. I don't have any theory-neutral argument for why biological and artifact functions are distinct. I treat their distinctness as an interesting consequence of the selected effects theory.

1.5 Rules of the Game

Before setting out to discover what functions are, we should lay down some ground rules for how to proceed in this intellectual endeavor. What is a "theory of function" supposed to be? How do we decide between competing theories of function? What kinds of evidence should we take seriously?

Let me go back to the beginning, but this time from a *meta* perspective. I started with a question: What are functions? By about halfway through the book, I will have defended an answer of the form, "functions are X." What is this statement supposed to be? Is it a report about how people think about functions? Or is it a report about what the world is like regardless of what anyone thinks, like "water is H_2O"? And if it's a report about how people think about functions, is it a report about how scientists think about them? Or is it about how ordinary people think about them? Or is it something like a recommendation for how we *ought* to think about them, regardless of what the world is like and regardless of what anyone happens to think now? Sometimes when philosophers seem to disagree about what functions are, they're really disagreeing about these meta questions. For example, I take it that when Ruse and Plantinga say that functions are intended effects, they mean to report something, first and foremost, about how ordinary people think about functions, about the concepts they have in their minds, and they might even be right – but that might still be irrelevant to my own project.

Here, I take the second approach: For me, a theory of function is a report about what exists in the world; it's not a report about what's in people's heads. Coming up with a theory of function is just like coming up with a theory about what gold is, or what aluminum is, or what spiders are. To know what aluminum is, you wouldn't, first of all, canvas people's opinions about it. You would get a chunk of it and start poking and prodding it.

https://doi.org/10.1017/9781108560764.002 Published online by Cambridge University Press

Other people might be more interested in understanding our concepts of function, and I applaud their endeavors, but that's not what I'm after.

There's an extra layer of complexity here, however, since you can't entirely separate these two projects – that is, of figuring out what functions are, and figuring out how people think about them. Sadly, functions aren't like aluminum or spiders. You can't just collect a sample and start poking at it. Functions are more abstract; they're less tangible. So you need a somewhat more abstract method to get at them. It seems to me that the very best way to figure out what functions are is to look at how biologists think about them. After all, if anyone knows what functions are, it's the biologists. For all that, I'm not trying to come up with a theory of what biologists think; I'm turning to biological usage to extract vital clues about what functions are. Looking at how biologists think about functions is a means to an end, not the end itself.

Now, on to the question of evidence. What kinds of evidence would prove that I'm right? There are three kinds of evidence I take seriously, ranked in terms of how seriously I take them: ordinary biological usage, ordinary biological practice, and bald intuitions. First and foremost, I think the best way to approach functions is to look at ordinary biological usage – that is, how biologists talk about them – as that talk is captured in sober scientific sources. That's what I did with Caro, and that led us to a discovery of the utmost importance: In ordinary biological usage, functions have explanatory depth. That doesn't mean scientists can't be wrong about particular cases, nor does it mean that scientists always use the word "function" in the same way.

Note that, even though I care about how biologists use "function," I'm not too interested in what biologists say functions are: that is, how biologists themselves would define the term "function" if you asked them, point blank, for a definition. "Function" is rarely defined in any explicit way in biology, and I wouldn't expect biologists to be able to state the rules that govern their use of "function." Just as you can use grammar flawlessly, without being able to state the rules of grammar, one can use "function" appropriately without being able to define it precisely.

Second, in some cases I consider the way functions are used in ordinary biological practice, particularly in biomedicine and biological psychiatry. Here's an example: When biomedical researchers say that a trait is *dysfunctional*, they're often indicating, in a pragmatic kind of way, that the trait is an appropriate target for medical intervention. It's the kind of thing you might want to fix or replace. This is a fact about biological practice: It's a fact about what sorts of actions biologists think are appropriate to take upon discovering that something has a function or that it's dysfunctional.

I put ordinary biological usage and ordinary biological practice above intuitions, since I think intuitions are mainly good for limning concepts; but they still have a role to play, when they're informed by good science. To the extent that I take intuitions seriously, I take them more seriously when they're about true-to-life cases (Do clay crystals have functions? What about piles of rocks?) rather than about science-fiction cases. (If the world were created five seconds ago in its present form, would anything have functions?) Sometimes people complain that intuitions have no role in serious philosophy – but soon enough, they start expressing their intuitions on all kinds of topics (they just don't *call* them intuitions). I might as well be up front about this fact from the outset.

Where have we reached so far? Functions are more complicated than they seem. The function of a trait is not just whatever it does. To illuminate functions, we considered the fact that sometimes, when biologists give functions to traits, they purport to explain why those traits exist. We then confronted this idea with the problem of backwards causation and considered whether there are any non-theist and nonfictionalist accounts of function that could solve the problem of backwards causation. I indicated that, if functions are selected effects, then we can understand their explanatory role quite easily. I then clarified what, precisely, the explanandum of the function statement is. Function statements purport to explain either why some entity has a trait, or why the members of a collection of entities generally have a trait. I recommended that we restrict functions to features (physical, psychological, or behavioral) of individuals or groups; individuals or groups themselves don't have functions. I urged that we not confuse biological and artifact functions, but the proof of that still awaits, and I set down some ground rules for the kinds of evidence that matter for defending a theory of functions.